IFWO

RAW SEQUENCE LISTING DATE: 08/30/2004
PATENT APPLICATION: **US/10/828,985A** TIME: 14:44:14

Input Set : **A:\RS0210Y.txt**
Output Set: **N:\CRF4\08302004\J828985A.raw**

<110> APPLICANT: Armour, Christopher D
Castle, John C
Garrett-Engele, Philip W
Kan, Zhengyan
Loerch, Patrick M
Tsinoremas, Nicholas F
<120> TITLE OF INVENTION: Novel Isoforms of Centromere Protein E (CENPE)
<130> FILE REFERENCE: RS0210Y
<140> CURRENT APPLICATION NUMBER: US 10/828,985A
<141> CURRENT FILING DATE: 2004-04-21
<150> PRIOR APPLICATION NUMBER: US 60/464,905
<151> PRIOR FILING DATE: 2003-04-23
<150> PRIOR APPLICATION NUMBER: US 60/510,701
<151> PRIOR FILING DATE: 2003-10-10
<160> NUMBER OF SEQ ID NOS: 25
<170> SOFTWARE: PatentIn version 3.2
<210> SEQ ID NO: 1
<211> LENGTH: 40
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 1
aagatgaatt acagaaaaag atccaagaac ttcagaaaaa 40
<210> SEQ ID NO: 2
<211> LENGTH: 40
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 2
taagggaaat gatagctaga gaccgacaga accaccaagt 40
<210> SEQ ID NO: 3
<211> LENGTH: 40
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 3
aagatgaatt acagaaaaag gaccgacaga accaccaagt 40
<210> SEQ ID NO: 4
<211> LENGTH: 40
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 4
aaactaaaaa agatcaagag aatgaactca gttcaaaagt 40
<210> SEQ ID NO: 5
<211> LENGTH: 40
<212> TYPE: DNA

RAW SEQUENCE LISTING DATE: 08/30/2004
PATENT APPLICATION: **US/10/828,985A** TIME: 14:44:14

Input Set : **A:\RS0210Y.txt**
Output Set: **N:\CRF4\08302004\J828985A.raw**

68 <213> ORGANISM: Homo sapiens
70 <400> SEQUENCE: 5
71 aaactaaaaa agatcaagag gaaagcattg aagacccaaa 40
74 <210> SEQ ID NO: 6
75 <211> LENGTH: 7704
76 <212> TYPE: DNA
77 <213> ORGANISM: Homo sapiens
79 <400> SEQUENCE: 6
80 atggcggagg aaggagccgt ggccgtctgc gtgcgagtgc ggccgctgaa cagcagagaa 60
82 gaatcacttg gagaaactgc ccaagtttac tggaaaactg acaataatgt catttatcaa 120
84 gttgatggaa gtaaatcctt caattttgat cgtgtctttc atggtaatga aactaccaaa 180
86 aatgtgtatg aagaaatagc agcaccaatc atcgattctg ccatacaagg ctacaatggt 240
88 actatatttg cctatggaca gactgcttca ggaaaaacat ataccatgat gggttcagaa 300
90 gatcatttgg gagttatacc cagggcaatt catgacattt tccaaaaaat taagaagttt 360
92 cctgataggg aatttctctt acgtgtatct tacatggaaa tatacaatga aaccattaca 420
94 gatttactct gtggcactca aaaaatgaaa cctttaatta ttcgagaaga tgtcaatagg 480
96 aatgtgtatg ttgctgatct cacagaagaa gttgtatata catcagaaat ggctttgaaa 540
98 tggattacaa agggagaaaa gagcaggcat tatggagaaa caaaaatgaa tcaaagaagc 600
100 agtcgttctc ataccatctt taggatgatt ttggaaagca gagagaaggg tgaaccttct 660
102 aattgtgaag gatctgttaa ggtatcccat ttgaatttgg ttgatcttgc aggcagtgaa 720
104 agagctgctc aaacaggcgc tgcaggtgtg cggctcaagg aaggctgtaa tataaatcga 780
106 agcttattta ttttgggaca agtgatcaag aaacttagtg atggacaagt tggtggtttc 840
108 ataaattatc gagatagcaa gttaacacga attctccaga attccttggg aggaaatgca 900
110 aagacacgta ttatctgcac aattactcca gtatcttttg atgaaacact tactgctctc 960
112 cagtttgcca gtactgctaa atatatgaag aatactcctt atgttaatga ggtatcaact 1020
114 gatgaagctc tcctgaaaag gtatagaaaa gaaataatgg atcttaaaaa acaattagag 1080
116 gaggtttctt tagagacgcg ggctcaggca atggaaaaag accaattggc ccaacttttg 1140
118 gaagaaaaag atttgcttca gaaagtacag aatgagaaaa ttgaaaactt aacacggatg 1200
120 ctggtgacct cttcttccct cacgttgcaa caggaattaa aggctaaaag aaaacgaaga 1260
122 gttacttggt gccttggcaa aattaacaaa atgaagaact caaactatgc agatcaattt 1320
124 aatataccaa caaatataac aacaaaaaca cataagcttt ctataaattt attacgagaa 1380
126 attgatgaat ctgtctgttc agagtctgat gttttcagta acactcttga tacattaagt 1440
128 gagatagaat ggaatccagc aacaaagcta ctaaatcagg agaatataga aagtgagttg 1500
130 aactcacttc gtgctgacta tgataatctg gtattagact atgaacaact acgaacagaa 1560
132 aaagaagaaa tggaattgaa attaaaagaa aagaatgatt tggatgaatt tgaggctcta 1620
134 gaaagaaaaa ctaaaaaaga tcaagagatg caactaattc atgaaatttc gaacttaaag 1680
136 aatttagtta agcatcgaga agtatataat caagatcttg agaatgaact cagttcaaaa 1740
138 gtagagctgc ttagagaaaa ggaagaccag attaagaagc tacaggaata catagactct 1800
140 caaaagctag aaaatataaa aatggacttg tcatactcat tggaaagcat tgaagaccca 1860
142 aaacaaatga agcagactct gtttgatgct gaaactgtag cccttgatgc caagagagaa 1920
144 tcagcctttc ttagaagtga aaatctggag ttgaaggaga aaatgaaaga acttgcaact 1980
146 acatacaagc aaatggaaaa tgatattcag ttatatcaaa gccaattgga ggcaaaaaag 2040
148 aaaatgcaag ttgatctgga gaaagaatta caatctgctt ttaatgagat aacaaaactc 2100
150 acctccctta tagatggcaa agttccaaaa gatttgctct gtaatttgga attggaagga 2160
152 aagattactg atcttcagaa agaactaaat aaagaagttg aagaaaatga agctttgcgg 2220
154 gaagaagtca ttttgctttc agaattgaaa tctttacctt ctgaagtaga aaggctgagg 2280
156 aaagagatac aagacaaatc tgaagagctc catataataa catcagaaaa agataaattg 2340
158 ttttctgaag tagttcataa ggagagtaga gttcaaggtt tacttgaaga aattgggaaa 2400
160 acaaaagatg acctagcaac tacacagtcg aattataaaa gcactgatca agaattccaa 2460

RAW SEQUENCE LISTING DATE: 08/30/2004
PATENT APPLICATION: **US/10/828,985A** TIME: 14:44:14

Input Set : **A:\RS0210Y.txt**
Output Set: **N:\CRF4\08302004\J828985A.raw**

```
162 aatttcaaaa cccttcatat ggactttgag caaaagtata agatggtcct tgaggagaat  2520
164 gagagaatga atcaggaaat agttaatctc tctaaagaag cccaaaaatt tgattcgagt  2580
166 ttgggtgctt tgaagaccga gctttcttac aagacccaag aacttcagga gaaaacacgt  2640
168 gaggttcaag aaagactaaa tgagatggaa cagctgaagg aacaattaga aaatagagat  2700
170 tctccgctgc aaactgtaga aagggagaaa acactgatta ctgagaaact gcagcaaact  2760
172 ttagaagaag taaaaacttt aactcaagaa aaagatgatc taaaacaact ccaagaaagc  2820
174 ttgcaaattg agagggacca actcaaaagt gatattcacg atactgttaa catgaatata  2880
176 gatactcaag aacaattacg aaatgctctt gagtctctga aacaacatca agaaacaatt  2940
178 aatacactaa aatcgaaaat ttctgaggaa gtttccagga atttgcatat ggaggaaaat  3000
180 acaggagaaa ctaaagatga atttcagcaa aagatggttg gcatagataa aaaacaggat  3060
182 ttggaagcta aaaataccca aacactaact gcagatgtta aggataatga gataattgag  3120
184 caacaaagga agatattttc tttaatacag gagaaaaatg aactccaaca aatgttagag  3180
186 agtgttatag cagaaaagga acaattgaag actgacctaa aggaaaatat tgaaatgacc  3240
188 attgaaaacc aggaagaatt aagacttctt ggggatgaac ttaaaaagca acaagagata  3300
190 gttgcacaag aaaagaacca tgccataaag aaagaaggag agctttctag gacctgtgac  3360
192 agactggcag aagttgaaga aaaactaaag gaaaagagcc agcaactcca agaaaaacag  3420
194 caacaacttc ttaatgtaca agaagagatg agtgagatgc agaaaaagat taatgaaata  3480
196 gagaatttaa agaatgaatt aaagaacaaa gaattgacat tggaacatat ggaaacagag  3540
198 aggcttgagt tggctcagaa acttaatgaa aattatgagg aagtgaaatc tataaccaaa  3600
200 gaaagaaaag ttctaaagga attacagaag tcatttgaaa cagagagaga ccaccttaga  3660
202 ggatatataa gagaaattga agctacaggc ctacaaacca aagaagaact aaaaattgct  3720
204 catattcacc taaaagaaca ccaagaaact attgatgaac taagaagaag cgtatctgag  3780
206 aagacagctc aaataataaa tactcaggac ttagaaaaat cccataccaa attacaagaa  3840
208 gagatcccag tgcttcatga ggaacaagag ttactgccta atgtgaaaaa agtcagtgag  3900
210 actcaggaaa caatgaatga actggagtta ttaacagaac agtccacaac caaggactca  3960
212 acaacactgg caagaataga aatggaaagg ctcaggttga atgaaaaatt tcaagaaagt  4020
214 caggaagaga taaaatctct aaccaaggaa agagacaacc ttaaaacgat aaaagaagcc  4080
216 cttgaagtta aacatgacca gctgaaagaa catattagag aaactttggc taaaatccag  4140
218 gagtctcaaa gcaaacaaga acagtcctta aatatgaaag aaaaagacaa tgaaactacc  4200
220 aaaatcgtga gtgagatgga gcaattcaaa cccaaagatt cagcactact aaggatagaa  4260
222 atagaaatgc tcggattgtc caaaagactt caagaaagtc atgatgaaat gaaatctgta  4320
224 gctaaggaga aagatgacct acagaggctg caagaagttc ttcaatctga aagtgaccag  4380
226 ctcaaagaaa acataaaaga aattgtagct aaacacctgg aaactgaaga ggaacttaaa  4440
228 gttgctcatt gttgcctgaa agaacaagag gaaactatta atgagttaag agtgaatctt  4500
230 tcagagaagg aaactgaaat atcaaccatt caaaagcagt tagaagcaat caatgataaa  4560
232 ttacagaaca agatccaaga gatttatgag aaagaggaac aacttaatat aaaacaaatt  4620
234 agtgaggttc aggaaaacgt gaatgaactg aaacaattca aggagcatcg caaagccaag  4680
236 gattcagcac tacaaagtat agaaagtaag atgctcgagt tgaccaacag acttcaagaa  4740
238 agtcaagaag aaatacaaat tatgattaag gaaaaagagg aaatgaaaag agtacaggag  4800
240 gcccttcaga tagagagaga ccaactgaaa gaaaacacta aagaaattgt agctaaaatg  4860
242 aaagaatctc aagaaaaaga atatcagttt cttaagatga cagctgtcaa tgagactcag  4920
244 gagaaaatgt gtgaaataga acacttgaag gagcaatttg agacccagaa gttaaacctg  4980
246 gaaaacatag aaacggagaa tataaggttg actcagatac tacatgaaaa ccttgaagaa  5040
248 atgagatctg taacaaaaga aagagatgac cttaggagtg tggaggagac tctcaaagta  5100
250 gagagagacc agctcaagga aaaccttaga gaaactataa ctagagacct agaaaaacaa  5160
252 gaggagctaa aaattgttca catgcatctg aaggagcacc aagaaactat tgataaacta  5220
254 agagggattg tttcagagaa aacaaatgaa atatcaaata tgcaaaagga cttagaacac  5280
256 tcaaatgatg ccttaaaagc acaggatctg aaaatacaag aggaactaag aattgctcac  5340
258 atgcatctga aagagcagca ggaaactatt gacaaactca gaggaattgt ttctgagaag  5400
```

RAW SEQUENCE LISTING DATE: 08/30/2004
PATENT APPLICATION: US/10/828,985A TIME: 14:44:14

Input Set : A:\RS0210Y.txt
Output Set: N:\CRF4\08302004\J828985A.raw

260 acagataaac tatcaaatat gcaaaaagat ttagaaaatt caaatgctaa attacaagaa 5460
262 aagattcaag aacttaaggc aaatgaacat caacttatta cgttaaaaaa agatgtcaat 5520
264 gagacacaga aaaaagtgtc tgaaatggag caactaaaga aacaaataaa agaccaaagc 5580
266 ttaactctga gtaaattaga aatagagaat ttaaatttgg ctcaagaact tcatgaaaac 5640
268 cttgaagaaa tgaaatctgt aatgaaagaa agagataatc taagaagagt agaggagaca 5700
270 ctcaaactgg agagagacca actcaaggaa agcctgcaag aaaccaaagc tagagatctg 5760
272 gaaatacaac aggaactaaa aactgctcgt atgctatcaa aagaacacaa agaaactgtt 5820
274 gataaactta gagaaaaaat ttcagaaaag acaattcaaa tttcagacat tcaaaaggat 5880
276 ttagataaat caaaagatga attacagaaa aaggaccgac agaaccacca agtaaaacct 5940
278 gaaaaaaggt tactaagtga tggacaacag caccttatgg aaagcctgag agaaaagtgc 6000
280 tctagaataa aagagctttt gaagagatac tcagagatgg atgatcatta tgagtgcttg 6060
282 aatagattgt ctcttgactt ggagaaggaa attgaattcc acagaatcat gaagaaactg 6120
284 aagtatgtgt taagctatgt tacaaaaata aaagaagaac aacatgaatg catcaataaa 6180
286 tttgaaatgg attttattga tgaagtggaa aagcaaaagg aattgctaat taaaatacag 6240
288 caccttcaac aagattgtga tgtaccatcc agagaattaa gggatctcaa attgaaccag 6300
290 aatatggatc tacatattga ggaaattctc aaagatttct cagaaagtga gttccctagc 6360
292 ataaagactg aatttcaaca agtactaagt aataggaaag aaatgacaca gtttttggaa 6420
294 gagtggttaa atactcgttt tgatatagaa aagcttaaaa atggcatcca gaaagaaaat 6480
296 gataggattt gtcaagtgaa taacttcttt aataacagaa taattgccat aatgaatgaa 6540
298 tcaacagagt ttgaggaaag aagtgctacc atatccaaag agtgggaaca ggacctgaaa 6600
300 tcactgaaag agaaaaatga aaaactattt aaaaactacc aaacattgaa gacttccttg 6660
302 gcatctggtg cccaggttaa tcctaccaca caagacaata agaatcctca tgttacatca 6720
304 agagctacac agttaaccac agagaaaatt cgagagctgg aaaattcact gcatgaagct 6780
306 aaagaaagtg ctatgcataa ggaaagcaag attataaaga tgcagaaaga acttgaggtg 6840
308 actaatgaca taatagcaaa acttcaagcc aaagttcatg aatcaaataa atgccttgaa 6900
310 aaaacaaaag agacaattca agtacttcag gacaaagttg ctttaggagc taagccatat 6960
312 aaagaagaaa ttgaagatct caaaatgaag cttgtgaaaa tagacctaga gaaaatgaaa 7020
314 aatgccaaag aatttgaaaa ggaaatcagt gctacaaaag ccactgtaga atatcaaaag 7080
316 gaagttataa ggctattgag agaaaatctc agaagaagtc aacaggccca agatacctca 7140
318 gtgatatcag aacatactga tcctcagcct tcaaataaac ccttaacttg tggaggtggc 7200
320 agcggcattg tacaaaacac aaaagctctt attttgaaaa gtgaacatat aaggctagaa 7260
322 aaagaaattt ctaagttaaa gcagcaaaat gaacagctaa taaaacaaaa gaatgaattg 7320
324 ttaagcaata atcagcatct ttccaatgag gtcaaaactt ggaaggaaag aacccttaaa 7380
326 agagaggctc acaaacaagt aacttgtgag aattctccaa agtctcctaa agtgactgga 7440
328 acagcttcta aaaagaaaca aattacaccc tctcaatgca aggaacggaa tttacaagat 7500
330 cctgtgccaa aggaatcacc aaaatcttgt ttttttgata gccgatcaaa gtctttacca 7560
332 tcacctcatc cagttcgcta ttttgataac tcaagtttag gcctttgtcc agaggtgcaa 7620
334 aatgcaggag cagagagtgt ggattctcag ccaggtcctt ggcacgcctc ctcaggcaag 7680
336 gatgtgcctg agtgcaaaac tcag 7704
339 <210> SEQ ID NO: 7
340 <211> LENGTH: 2568
341 <212> TYPE: PRT
342 <213> ORGANISM: Homo sapiens
344 <400> SEQUENCE: 7
346 Met Ala Glu Glu Gly Ala Val Ala Val Cys Val Arg Val Arg Pro Leu
347 1 5 10 15
350 Asn Ser Arg Glu Glu Ser Leu Gly Glu Thr Ala Gln Val Tyr Trp Lys
351 20 25 30
354 Thr Asp Asn Asn Val Ile Tyr Gln Val Asp Gly Ser Lys Ser Phe Asn

RAW SEQUENCE LISTING DATE: 08/30/2004
PATENT APPLICATION: **US/10/828,985A** TIME: 14:44:14

Input Set : **A:\RS0210Y.txt**
Output Set: **N:\CRF4\08302004\J828985A.raw**

```
355         35                  40                  45
358 Phe Asp Arg Val Phe His Gly Asn Glu Thr Thr Lys Asn Val Tyr Glu
359     50                  55                  60
362 Glu Ile Ala Ala Pro Ile Ile Asp Ser Ala Ile Gln Gly Tyr Asn Gly
363 65                  70                  75                  80
366 Thr Ile Phe Ala Tyr Gly Gln Thr Ala Ser Gly Lys Thr Tyr Thr Met
367                 85                  90                  95
370 Met Gly Ser Glu Asp His Leu Gly Val Ile Pro Arg Ala Ile His Asp
371             100                 105                 110
374 Ile Phe Gln Lys Ile Lys Lys Phe Pro Asp Arg Glu Phe Leu Leu Arg
375         115                 120                 125
378 Val Ser Tyr Met Glu Ile Tyr Asn Glu Thr Ile Thr Asp Leu Leu Cys
379     130                 135                 140
382 Gly Thr Gln Lys Met Lys Pro Leu Ile Ile Arg Glu Asp Val Asn Arg
383 145                 150                 155                 160
386 Asn Val Tyr Val Ala Asp Leu Thr Glu Glu Val Val Tyr Thr Ser Glu
387                 165                 170                 175
390 Met Ala Leu Lys Trp Ile Thr Lys Gly Glu Lys Ser Arg His Tyr Gly
391             180                 185                 190
394 Glu Thr Lys Met Asn Gln Arg Ser Ser Arg Ser His Thr Ile Phe Arg
395         195                 200                 205
398 Met Ile Leu Glu Ser Arg Glu Lys Gly Glu Pro Ser Asn Cys Glu Gly
399     210                 215                 220
402 Ser Val Lys Val Ser His Leu Asn Leu Val Asp Leu Ala Gly Ser Glu
403 225                 230                 235                 240
406 Arg Ala Ala Gln Thr Gly Ala Ala Gly Val Arg Leu Lys Glu Gly Cys
407                 245                 250                 255
410 Asn Ile Asn Arg Ser Leu Phe Ile Leu Gly Gln Val Ile Lys Lys Leu
411             260                 265                 270
414 Ser Asp Gly Gln Val Gly Gly Phe Ile Asn Tyr Arg Asp Ser Lys Leu
415         275                 280                 285
418 Thr Arg Ile Leu Gln Asn Ser Leu Gly Gly Asn Ala Lys Thr Arg Ile
419     290                 295                 300
423 Ile Cys Thr Ile Thr Pro Val Ser Phe Asp Glu Thr Leu Thr Ala Leu
424 305                 310                 315                 320
427 Gln Phe Ala Ser Thr Ala Lys Tyr Met Lys Asn Thr Pro Tyr Val Asn
428                 325                 330                 335
431 Glu Val Ser Thr Asp Glu Ala Leu Leu Lys Arg Tyr Arg Lys Glu Ile
432             340                 345                 350
435 Met Asp Leu Lys Lys Gln Leu Glu Glu Val Ser Leu Glu Thr Arg Ala
436         355                 360                 365
439 Gln Ala Met Glu Lys Asp Gln Leu Ala Gln Leu Leu Glu Glu Lys Asp
440     370                 375                 380
443 Leu Leu Gln Lys Val Gln Asn Glu Lys Ile Glu Asn Leu Thr Arg Met
444 385                 390                 395                 400
447 Leu Val Thr Ser Ser Ser Leu Thr Leu Gln Gln Glu Leu Lys Ala Lys
448                 405                 410                 415
451 Arg Lys Arg Arg Val Thr Trp Cys Leu Gly Lys Ile Asn Lys Met Lys
453             420                 425                 430
```

VERIFICATION SUMMARY DATE: 08/30/2004
PATENT APPLICATION: **US/10/828,985A** TIME: 14:44:15

Input Set : **A:\RS0210Y.txt**
Output Set: **N:\CRF4\08302004\J828985A.raw**